

(19) World Intellectual Property Organization  
International Bureau



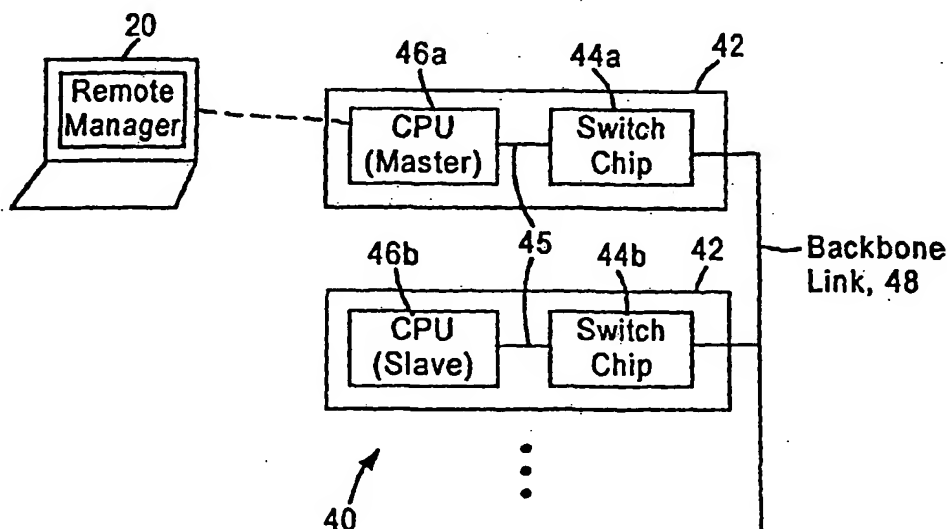
(43) International Publication Date  
24 January 2002 (24.01.2002)

PCT

(10) International Publication Number  
**WO 02/07383 A2**

- (51) International Patent Classification<sup>7</sup>: **H04L 12/00**
- (21) International Application Number: **PCT/US01/09539**
- (22) International Filing Date: **23 March 2001 (23.03.2001)**
- (25) Filing Language: **English**
- (26) Publication Language: **English**
- (30) Priority Data:  
**09/618,057**      **17 July 2000 (17.07.2000)**      **US**
- (71) Applicant: **ADVANCED MICRO DEVICES, INC.**  
[US/US]; Mail Stop 68, One AMD Place, Sunnyvale, CA  
94088-3453 (US).
- (72) Inventor: **EGBERT, Chandan**; 3632 Bloomsbury Way,  
San Jose, CA 95132 (US).
- (74) Agent: **RODDY, Richard, J.**; Advanced Micro Devices,  
Inc., Mail Stop 68, One AMD Place, Sunnyvale, CA 94088-  
3453 (US).
- (81) Designated States (*national*): **BR, CN, JP, KR, SG.**
- (84) Designated States (*regional*): **European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE, TR).**
- Published:**  
— *without international search report and to be republished  
upon receipt of that report*
- For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.*

(54) Title: **IN-BAND MANAGEMENT OF A STACKED GROUP OF SWITCHES BY A SINGLE CPU**



(57) Abstract: A switching system includes switches, each having a host processing unit and a switching unit, and a backbone link configured for transferring data packets between the switching units. One of the host processing units is configured as a master unit for generating a data frame having a destination address for a selected one of the switching units of a corresponding selected one of the other host processing units. The master unit outputs the data frame to the corresponding switching unit for transfer to the selected one switching unit via the backbone link. The selected one switching unit, in response to receiving the data frame having the corresponding destination address, forwards the data frame to the corresponding host processing unit for execution of a processing operation specified in the data frame. Hence, the switching system provides inter-processor communications using a preexisting backbone link, eliminating the necessity of a processor bus.

## IN-BAND MANAGEMENT OF A STACKED GROUP OF SWITCHES BY A SINGLE CPU

### FIELD OF THE INVENTION

The present invention relates to management of a stack of layer 2 switches configured for switching data packets.

### BACKGROUND ART

5 Local area networks use a network cable or other media to link stations on the network. Each local area network architecture uses a media access control (MAC) enabling network interface devices at each network node to access the network medium.

The Ethernet protocol IEEE 802.3 has evolved to specify a half-duplex media access mechanism and a full-duplex media access mechanism for transmission of data packets. The full-  
10 duplex media access mechanism provides a two-way, point-to-point communication link between two network elements, for example between a network node and a switched hub.

Switched local area networks are encountering increasing demands for higher speed connectivity, more flexible switching performance, and the ability to accommodate more complex network architectures. For example, commonly-assigned U.S. Patent No. 5,953,335 discloses a  
15 network switch configured for switching layer 2 type Ethernet (IEEE 802.3) data packets between different network nodes; a received data packet may include a VLAN (virtual LAN) tagged frame according to IEEE 802.1q protocol that specifies another subnetwork (via a router) or a prescribed group of stations. The network switch includes multiple switch ports configured for sending and receiving data packets to respective network nodes according to IEEE 802.3 protocol, plus a  
20 management port that enables an external management entity to control overall operations of the switch using a management MAC interface. Hence, a host controller such as a host CPU can access the network switch chip via a standard MAC layer protocol.

Figure 1 is a diagram illustrating a switch rack 10 having a stacked group of switches 12 for deployment of a local area network, managed by a remote management station 20. Each switch 12 has  
25 an integrated multiport switch 14 as disclosed in U.S. Patent No. 5,953,355, plus a corresponding host CPU 16 for controlling the corresponding integrated multiport switch 14. Remote management software executed by the remote management station 20 manages the switches 12 by instructing a master CPU to read and write control/status registers in the switch chips 14. The switch rack 10 also includes a CPU bus 18 configured for communications between the CPUs 16, and an backbone link 22  
30 for transfer of data frames between the integrated multiport switches 14. In particular, one of the CPUs 16a is designated as the master CPU and the remaining CPUs (e.g., 16b) are designated as slave

CPUs, eliminating the necessity for the remote management station 20 to communicate with each of the individual CPUs 16. Hence, the remote management station 20 need only communicate with the master CPU 16a, with the master CPU 16a controlling the remaining slave CPUs (e.g., writing information into CPU control/status registers) via the CPU bus 18.

- 5        Use of the CPU bus 18, however, increases the required pin count of the CPUs 16, and the complexity of the switch rack 10.

### SUMMARY OF THE INVENTION

10        There is a need for an arrangement that enables communication between multiple host processing units, configured for controlling respective network switch devices, without the necessity of a separate CPU bus.

15        These and other needs are attained by the present invention, where a switching system includes switches, each having a host processing unit and a switching unit, and a backbone link configured for transferring data packets between the switching units. One of the host processing units is configured as a master unit for generating a data frame having a destination address for a selected one of the switching units of a corresponding selected one of the other host processing units. The master unit outputs the data frame to the corresponding switching unit for transfer to the selected one switching unit via the backbone link. The selected one switching unit, in response to receiving the data frame having the corresponding destination address, forwards the data frame to the corresponding host  
20        processing unit for execution of a processing operation specified in the data frame. Hence, the switching system provides inter-processor communications using a preexisting backbone link, eliminating the necessity of a processor bus.

25        One aspect of the present invention provides a method in a network switching system having multiple switching units controlled by respective host processing units. The method includes generating a data frame by a first of the host processing units, configured as a master unit, by including a destination address of a selected one of the switching units controlled by a corresponding selected one of the other host processing units. The method also includes outputting the data frame from the master unit to the corresponding switching unit, transferring the data frame from the switching unit having received the data frame from the master unit to the selected one switching unit via a backbone  
30        link, and forwarding the data frame from the selected one switching unit to the corresponding selected one other host processing unit, for execution of a processing operation specified in the data frame, based on the selected one switching unit detecting the destination address matching an assigned address of the selected one switching unit.

35        Another aspect of the present invention provides a switching system configured for receiving commands from a remote manager. The system includes switches and a backbone link. Each switch

has a switching unit, having an assigned network address, and a corresponding host processing unit configured for controlling the corresponding switching unit. One of the host processing units is configured for receiving a command from the remote manager and in response generating a data frame having a destination address for configuration of a selected one of the switching units by the  
5 corresponding selected one of the other host processing units. Each switching unit is configured for forwarding a received data frame having a matching destination address to the corresponding host processing unit. The backbone link is configured for transferring the data frame from the switching unit corresponding to the one host processing unit to the selected switching unit, and the selected one other host processing unit is configured for executing a processing operation specified in the data  
10 frame based on retrieval of the data frame from the corresponding selected switching unit.

Additional advantages and novel features of the invention will be set forth in part in the description which follows and in part will become apparent to those skilled in the art upon examination of the following or may be learned by practice of the invention. The advantages of the present invention may be realized and attained by means of instrumentalities and combinations particularly  
15 pointed in the appended claims.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Reference is made to the attached drawings, wherein elements having the same reference numeral designations represent like elements throughout and wherein:

Figure 1 is a diagram illustrating a conventional (prior art) stacked rack of managed switches.

20 Figure 2 is a block diagram of a stacked rack of managed switches according to an embodiment of the present invention.

Figure 3 is a diagram illustrating the method of managing multiple switching units using a master host processing unit according to an embodiment of the present invention.

#### BEST MODE FOR CARRYING OUT THE INVENTION

25 Figure 2 is a block diagram illustrating a switch rack 40 having a stacked group of switches 42 for a packet switched network, such as an Ethernet (IEEE 802.3) network. Each switch 42 includes an integrated (i.e., single chip) multiport switch 44, also referred to as a switching unit, similar to the switch described in U.S. Patent No. 5,953,355, that enables communication of data packets between network stations, for example a client workstation, at 10 Mbps or 100 Mbps according to IEEE 802.3 protocol.  
30 Each switch 42 also includes a corresponding host processing unit 46 configured for controlling the corresponding switching unit 44 via a link 45, for example a PCI local bus or a MAC (IEEE 802.3) based data link. Each of the integrated multiport switches 44 are interconnected by a backbone link 48, implemented for example as a gigabit Ethernet link coupled to expansion ports of the switch chips 44, enabling transfer of data packets between subnetworks served by the respective switch chips 44.

One of the host processing units, for example unit 46a, is configured as a master unit for reception of commands from the remote manager 20. In particular, the remote manager 20 sends instructions to the master CPU 46a to perform processing operations, for example reading and writing information from and to control/status registers in the switch chips 44. The master CPU 46a in turn sends data frames to the remaining host processing units (i.e., slave CPUs), such as processing unit 46b, instructing the slave CPUs to perform the same processing operations requested by the remote manager 20.

Figure 3 is a flow diagram illustrating the method of sending commands to slave CPUs 46 via the backbone link 48 according to an embodiment of the present invention. As described below, the sending of commands to the slave CPUs via the backbone link 48 eliminates the necessity of the CPU bus of Figure 1.

The method begins in step 50, where the master CPU 46a receives a command from the remote manager 20. A command from the remote manager 20 may specify greeting or writing data to a prescribed control/status register in the switching unit 44. The master host processing unit 46a responds to the command from the remote manager 20 by performing the appropriate processing operation within the corresponding switching unit 44a in step 52. The master host processing unit 46a then generates a data frame for each of the slave processing units (e.g., 46b) by inserting the destination address for the corresponding selected switching unit 44 in step 54, by adding a prescribed value specifying the processing operation within the Ethertype field in step 56, and adding the remaining data for execution of the processing operation within the payload in step 58. In particular, the Ethertype field is a 16-bit value, where certain values are unused in Ethernet protocol; accordingly, the unused values are utilized to specify prescribed operations for the slave CPUs.

The master host processing unit 46a outputs each of the data frames to the corresponding switching unit 44a via the link 45 in step 60, for example as a management frame or as a PCI transfer. The switching unit 44a, recognizing from its internal frame forwarding tables that the destination addresses are reachable a via the backbone link 48, outputs the data frames in step 62 onto the backbone link 48.

The switching units (e.g., 44b) corresponding to the slave CPUs receive the data frames in step 64, and perform a lookup of the destination address in order to make a frame forwarding decision. Each slave switching unit 44 forwards the received data frame to the corresponding host processing unit 46 via the corresponding link 45 in step 68 based on a determined match between the destination MAC address and the assigned MAC address of the switching unit 44. Hence, each slave CPU (e.g., 46b) is able to receive the corresponding data frame from the master CPU 46a based on a match between the destination address in the received data frame and the assigned MAC address of the corresponding switching unit 44.

The slave CPU 46b, having received the data frame, parses the data frame in step 70 and determines from the Ethertype field and the payload the processing operation to be performed. The slave CPU 46b performs the specified processing operation (e.g., reading or writing to a control/status register in

the corresponding switching unit 44) in step 72, and generates in step 74 a reply data frame for the master CPU 46a acknowledging execution of the requested processing operation.

While this invention has been described with what is presently considered to be the most practical preferred embodiment, it is to be understood that the invention is not limited to the disclosed  
5 embodiments, but, on the contrary, is intended to cover various modifications and equivalent arrangements included within the spirit and scope of the appended claims.

What is Claimed Is:

1. A method in a network switching system having multiple switching units controlled by respective host processing units, the method comprising:  
generating a data frame by a first of the host processing units, configured as a master unit, by including a destination address of a selected one of the switching units controlled by a corresponding  
5 selected one of the other host processing units;  
outputting the data frame from the master unit to the corresponding switching unit;  
transferring the data frame from the switching unit having received the data frame from the master unit to the selected one switching unit via a backbone link; and  
forwarding the data frame from the selected one switching unit to the corresponding selected  
10 one other host processing unit, for execution of a processing operation specified in the data frame, based on the selected one switching unit detecting the destination address matching an assigned address of the selected one switching unit.
2. The method of claim 1, wherein the generating step includes inserting a prescribed value specifying at least a portion of the processing operation in an Ethernet field within the data frame.
3. The method of claim 2, wherein the transferring step includes outputting the data frame onto the backbone link according to IEEE 802.3 protocol.
4. The method of claim 2, wherein the outputting step includes outputting the data frame from the master unit to the corresponding switching unit according to IEEE 802.3 protocol.
5. The method of claim 2, wherein the outputting step includes outputting the data frame from the master unit to the corresponding switching unit using a peripheral component interconnect (PCI) bus transfer.
6. The method of claim 1, wherein the generating step includes generating the data frame based on a command received from a remote manager.
7. The method of claim 6, wherein the generating step further includes configuring by the master unit each of the other host processing units, based on the command received from the remote

manager, by generating a corresponding data frame having a destination address specifying the corresponding switching unit.

8. The method of claim 6, further comprising configuring the selected one switching unit by the corresponding selected one other host processing unit based on parsing the received data frame.

9. The method of claim 8, wherein the configuring step includes writing prescribed values, recovered from the received data frame, into control/status registers within the selected one switching unit specified by the received data frame.

10. A switching system configured for receiving commands from a remote manager, the system comprising:

- switches, each having a switching unit, having an assigned network address, and a corresponding host processing unit configured for controlling the corresponding switching unit, one of the host processing units configured for receiving a command from the remote manager and in response generating a data frame having a destination address for configuration of a selected one of the switching units by the corresponding selected one of the other host processing units, each switching unit configured for forwarding a received data frame having a matching destination address to the corresponding host processing unit; and
- 10 a backbone link configured for transferring the data frame from the switching unit corresponding to the one host processing unit to the selected switching unit, the selected one other host processing unit configured for executing a processing operation specified in the data frame based on retrieval of the data frame from the corresponding selected switching unit.

11. The system of claim 10, wherein the one host processing unit inserts a value within an Ethertype field of the data frame that specifies at least a portion of the processing operation.

12. The system of claim 11, wherein the backbone link is configured for transferring the data frame according to IEEE 802.3 protocol.

13. The switch of claim 10, wherein the selected one other host processing unit is configured for writing prescribed values, recovered from the received data frame, into control/status registers within the selected one switching unit specified by the received data frame.



14. The switch of claim 10, wherein the selected one other host processing unit is configured for generating a second data frame having the destination address of the switching unit corresponding to the one host processing unit and specifying results based on the processing operation.

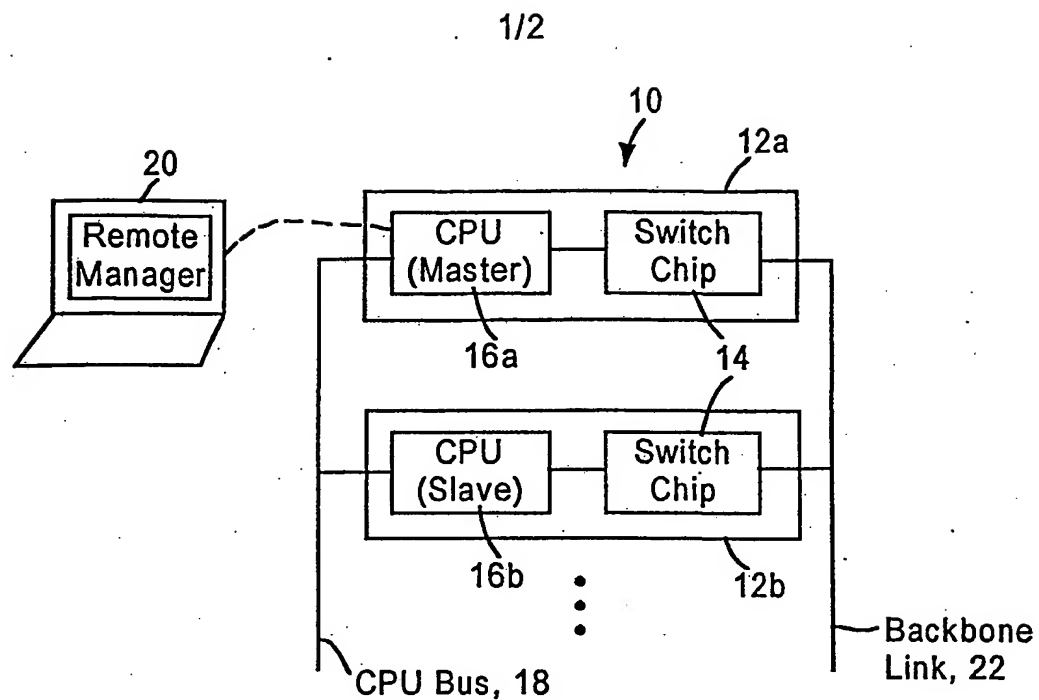


FIG. 1  
(PRIOR ART)

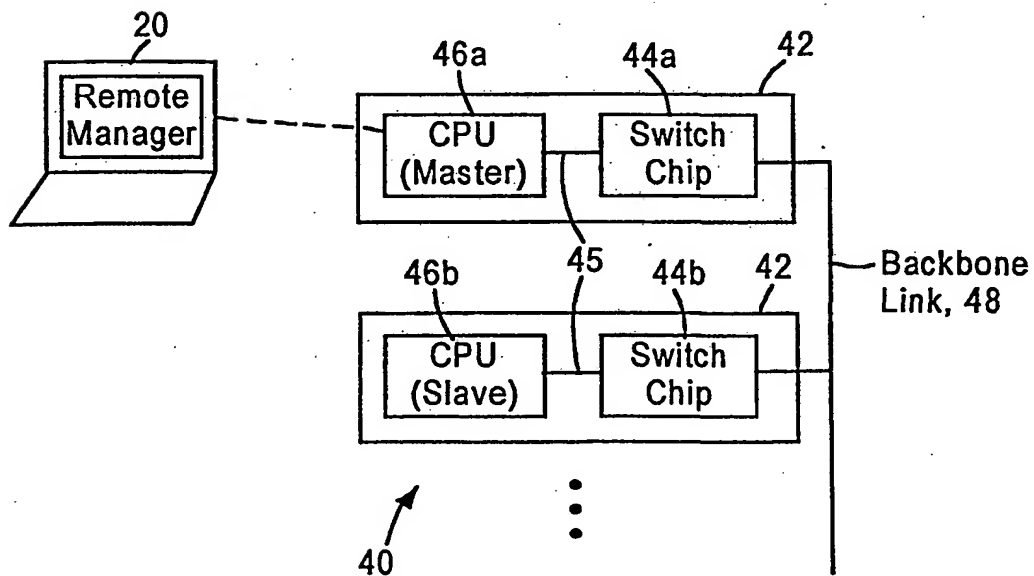


FIG. 2

2/2

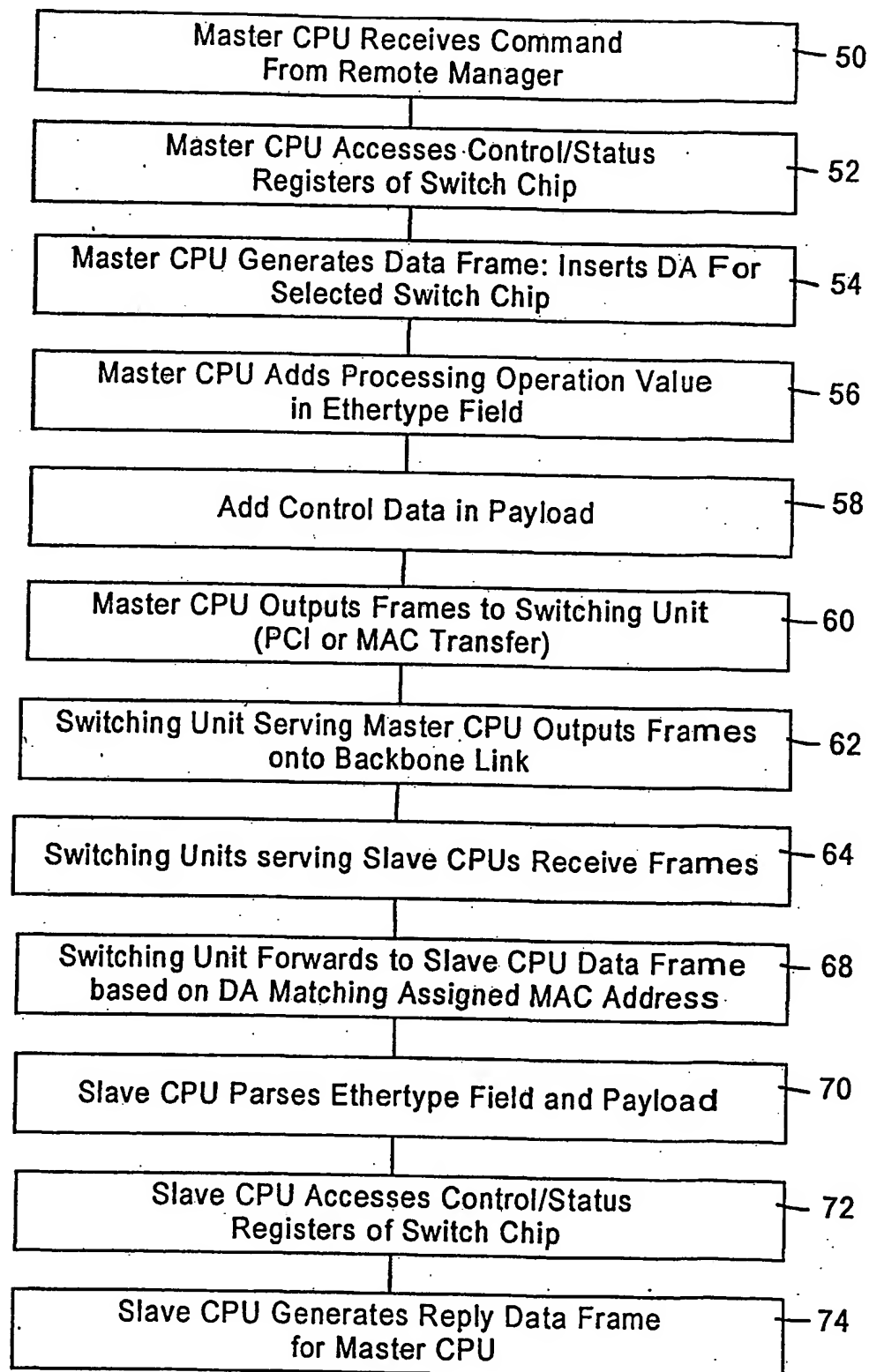


FIG. 3

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
24 January 2002 (24.01.2002)

PCT

(10) International Publication Number  
WO 02/07383 A3

(51) International Patent Classification<sup>7</sup>: H04L 12/403,  
12/44, 12/24

(74) Agent: RODDY, Richard, J.: Advanced Micro Devices,  
Inc., Mail Stop 68, One AMD Place, Sunnyvale, CA 94088-  
3453 (US).

(21) International Application Number: PCT/US01/09539

(81) Designated States (*national*): BR, CN, JP, KR, SG.

(22) International Filing Date: 23 March 2001 (23.03.2001)

(84) Designated States (*regional*): European patent (AT, BE,  
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,  
NL, PT, SE, TR).

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
09/618,057 17 July 2000 (17.07.2000) US

Published:  
— with international search report

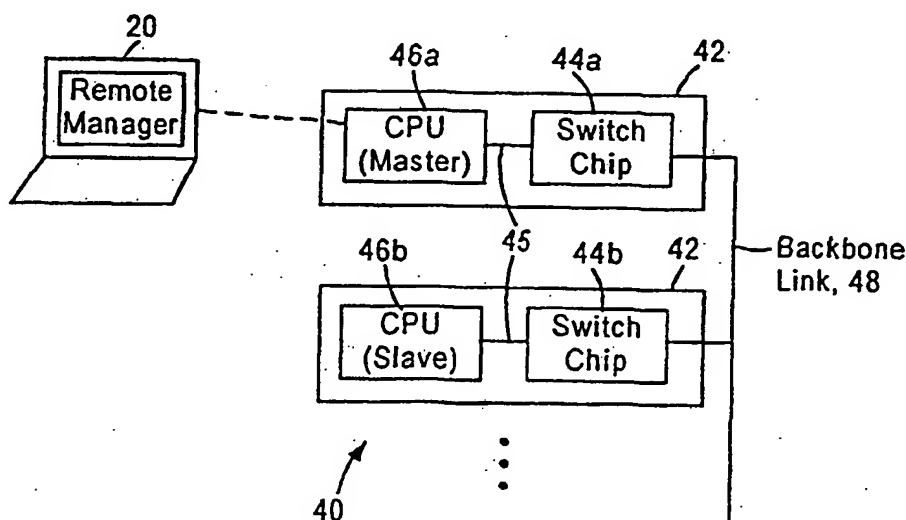
(71) Applicant: ADVANCED MICRO DEVICES, INC.  
[US/US]: Mail Stop 68, One AMD Place, Sunnyvale, CA  
94088-3453 (US).

(88) Date of publication of the international search report:  
16 May 2002

(72) Inventor: EGBERT, Chandan: 3632 Bloomsbury Way,  
San Jose, CA 95132 (US).

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: IN-BAND MANAGEMENT OF A STACKED GROUP OF SWITCHES BY A SINGLE CPU



(57) Abstract: A switching system includes switches, each having a host processing unit and a switching unit, and a backbone link configured for transferring data packets between the switching units. One of the host processing units is configured as a master unit for generating a data frame having a destination address for a selected one of the switching units of a corresponding selected one of the other host processing units. The master unit outputs the data frame to the corresponding switching unit for transfer to the selected one switching unit via the backbone link. The selected one switching unit, in response to receiving the data frame having the corresponding destination address, forwards the data frame to the corresponding host processing unit for execution of a processing operation specified in the data frame. Hence, the switching system provides inter-processor communications using a preexisting backbone link, eliminating the necessity of a processor bus.

WO 02/07383 A3

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 01/09539

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 H04L12/403 H04L12/44 H04L12/24

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, PAJ

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>US 6 084 878 A (CHOW PETER KA-FA ET AL) 4 July 2000 (2000-07-04) abstract column 1, line 25 - line 64 column 2, line 16 - line 51 column 3, line 21 - line 37 column 4, line 61 - column 5, line 2 claim 1</p> <p style="text-align: center;">-/-</p>	1-14

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*G\* document member of the same patent family

Date of the actual completion of the international search

24 January 2002

Date of mailing of the international search report

31/01/2002

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Lai, C

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 01/09539

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	<p>"ARCHITECTURE FOR STACKABLE HUB MANAGEMENT" IBM TECHNICAL DISCLOSURE BULLETIN, IBM CORP. NEW YORK, US, vol. 39, no. 2, 1 February 1996 (1996-02-01), pages 3-9, XP000559801 ISSN: 0018-8689 page 3, line 8 - line 16 page 5, line 12 - line 20 figure 5</p>	1-14

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 01/09539

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 6084878	A	04-07-2000	NONE